

SEQUENCE LISTING

<110> Tim Keith

<120> NOVEL HUMAN GENES RELATING TO RESPIRATORY DISEASES AND OBESITY

<130> HUM00-23

<141> 2000/07/28

<150> 60/211,749

<151> 2000/06/14

<150> 60/146,336

<151> 1999/07/30

<160> 11

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 10304

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(10304)

<223> n = A,T,C or G

<400> 1

cgggcggtgta	tatctcttca	tagagagcgc	tcagacagcg	tgcgttaatc	tgcgtcgata	60
tatagagatc	tttatcactg	agtagataga	acgtacatga	atgtacgaac	agtcacagcg	120
agtaacttga	ctaggataag	atagacagta	ccaactaatg	agacaagaag	agggaaatcat	180
atagaatcat	gtagtctgag	tctagcgagt	gtcgacatga	tcacaagcga	aatacagact	240
atgagaagag	gtagaaataa	taagtanact	gagaagagag	gtcatatgta	catacaaatac	300
agtaaagcaa	tagaaattga	atacattata	agccacagtt	acagaattag	cctaatttaa	360
caacctatggc	aagcgagtta	tatcaaacat	agaagagtaa	actctatcga	ccatgggtag	420
gaacgaataa	aggcgctcgag	aagacaataa	gaatgcgtgt	taaacagcaa	tacaagagaa	480
tagcaccact	gaagcagacc	aaaggcgctca	ccggggaagt	agggaagagg	cacctcacia	540
ggagaggaaa	gggcagtcct	gattttgaaa	atttcagtga	aaagacagtg	ttgttcccgg	600
aggcagctta	gtgatcccg	atcgactctg	aagaggaccc	tgagggtagg	ggatttttgg	660
gcctgaccgg	cctatgctga	acgcccaccg	ggaattcagg	gagaaacacg	gggccccggc	720
ttccaggaga	gcagccaggc	cacagccctg	aggacgggca	aacccccacc	aggcacgggtg	780
agagggaggc	cgcccaggcc	tggggcctgg	cggcagggga	tgaagtggac	cagagccccg	840
caaatacctaa	cgtgggtgag	cagtgagcct	gtgtggctgc	gagtggctcc	gttttggggc	900
tgttttgttcc	tgcagcaaat	gatgccagcc	ctgacggaac	cagtgcacgt	ccaccacgag	960
ctgcccacgt	cctctccagg	aagggaacccg	ggtccacgag	ctgcccacgt	cctctccagg	1020
aagggaacga	agaccacgag	ctgcccacgc	cctctccagg	aggggaacac	gggttcacga	1080
gctgcccacg	tcctctccag	gaggggacac	cgggttcagt	agctgcccac	gccctttcca	1140
ggaagggaac	ccgggttcac	gagctgcccc	cgtcctctcc	aggaggggac	accgggttca	1200
cgagctgccc	acgtcctctc	caggaaggga	cccagggtcca	cgaactgccc	acgcccctctc	1260
caggagggga	ccgggttcca	cgagctgccc	acgtcgtctc	caggaaggga	cccgggtcca	1320
cgagctgccc	acgtcctctc	caggaaggga	cccgggtcca	cgagctgccc	acgtcctctc	1380
caggaaggga	ccccgggttc	acgagctgcc	cacgtcctct	ccaggaaggg	accccgggtt	1440
cacgagctgc	ccacgtcctc	tccaggaggg	gaccccgggt	tcacgagctg	cccacgtcct	1500

ctccaggaag	ggacccccggg	ttcacgagct	gccacgctcg	tctccaggaa	gggaccccggg	1560
tccacgagct	gcccacgtcc	tctccaggaa	aggaccccggg	tccacgagct	ggcccacgtcc	1620
tctgcaggaa	gggacccccgg	gtccacgagc	tgcccacgtc	ctctccagga	agggacc:ccg	1680
ggttcacgag	ctgcccacgt	cctctccagg	aagggacccc	gggtccacga	gctgccc:acg	1740
tcctctccag	gaagggaacc	cggttccacg	aactgcccac	gtcctctcca	ggaagggiacc	1800
ccgggttcac	gagctgccc	cgctctctcc	aggaggggac	accgggttca	cgagctgccc	1860
acgccctctc	caggaaggga	ccccgggttc	atgagctgcc	cacgtcctct	ccaggaagg	1920
acccgggtcc	acgaactgcc	cacgcccctc	ccaggagggg	acccgggtcc	acgagctgcc	1980
cacgtcgtca	acgggaagg	acccgggtcc	acgagctgcc	cacgtcctct	ccaggaagg	2040
acccgggtcc	acgaactgcc	cacgcgctct	ccaggagggg	acacccgggt	cacgagctgc	2100
ccacgccctc	tccaggaagg	gacccccggg	tccacgagctg	cccacgtcct	ctccaggagg	2160
ggacaccggg	ttcacgagct	gcccacgtcc	tctccaggag	gggacaccgg	gttcacgagc	2220
tgcccacgcc	ctctccagga	ggggacaccg	ggttcacgag	ctgcccacgt	cctctccagg	2280
aagggaaccg	ggtccacgag	ctgcccacgt	cctctccagg	aggggacacc	gggttcacga	2340
gctgcccacg	cactttccag	gaagggaacc	cggttccagg	tctcctgccg	gcccacatcg	2400
tgctttgtg	taaatcagaa	gaaagatgag	gaacaggccc	tcctctctct	ccaggcaggc	2460
tttggtggag	gggctggatc	tctgcccga	ccttccctgg	cagggcacc	tgtgcttgag	2520
ccccagaact	gcaggcggcc	ggcagagaag	gggtccatga	tggcgccctg	gtgcgcagcc	2580
ttggacctgc	ccccatggac	ctgggtgagg	acttcccagc	ccttccccgg	ctccagctgc	2640
tctccctaag	ccgcctcacc	ccttccctcg	gcagggggca	gtggacgagg	gttccgtccc	2700
tccaggggat	gtcccaaac	ccctgccagg	acttggcaga	tccggcctct	catcttgga	2760
gctagatggt	gggacgggat	catcgtggtg	gctttaattt	gcatttctct	gatgac:gat	2820
gatttcgagc	atctcttcat	atgtttgctg	gctttgggga	tagagatatt	tcttccc:aaa	2880
gcaaaaactg	attatgtcat	ttctgcttca	agatgccagt	gatgcctgag	gtctgcagg	2940
cagtgcatac	gctcaccgcc	tgcccgctca	ggagcctgtg	cttgaccccc	aaatccgcc	3000
cccaactccc	tggtaccggc	tactccttc	catgaggggc	ccttccccagg	gacagccgat	3060
gctctcctga	tggctcctgc	ccttgccagg	tgtgcccccc	gcctgcccac	ctggcctgga	3120
ccctcgccctg	agccccctca	gggtctctcg	ccacctcaac	ccaggcggtt	gttccgcagg	3180
aacctcccgg	ctcttcccac	tccggaaaag	aaggctctgg	gcattggagg	cggcagggcc	3240
ccatccccgt	accctggccc	ttcttccctg	tctctggttg	tccctgcccc	ggggcccttg	3300
cacctgcatt	ccctctctct	gtgagtgccc	tggggcccgt	tacccacgtc	accgtcccag	3360
gatacctttt	cttttcttcc	tctctctcca	gctttattga	ggtatagtgt	acaattcagg	3420
acggtgtgca	ctcaagggtat	gcagcatcac	aacctgacac	acgtaggcat	tgtgaaatga	3480
gtcccacaat	tgggctaatt	aacacaccca	tcaccttaca	tggttacttc	tttctgtggt	3540
gagaacacta	aatttttaaat	agaggacaca	cagcctgggc	aacatagtga	gaccctgtct	3600
ctacaaatat	aaaaaaatta	tctggacgtg	gtggtgcaca	cctgtggtcc	cagctacttg	3660
ggaagctgag	gctggagaat	cacttgagcc	tgggaggcgg	aggttgccgt	gcactccagc	3720
ctgggcgaca	gagggaggcc	ctatctcaaa	ataaataaat	aaaggacaca	ttcttatcag	3780
ctgtagtca	cacgttcatt	acatcttaga	acccgctaatt	ctcataactg	cacctttggt	3840
ccctgtgacc	ctcaactccc	ggtcccctcc	agccctgaca	gccactgttc	actctgcttc	3900
tgtgagttcc	gctttttcac	acgtcactcg	agtgaggcca	tgtgctgttt	gtctttctgt	3960
gcctggetta	tctcacttac	cacaaatgcc	ccttcagggtc	atcgtgtcct	cacaaatggc	4020
gggcttgccc	tgccctgccc	tgccctgccc	tcccttccct	tcccttctct	ctctctcctt	4080
tctctctctc	tggctctctc	tctctcccac	ccttcccttt	ccctcctgtg	gaataaact	4140
cctgtgtgtg	tgtgcatgca	tgtgtgtgta	tatttctcac	atattttcat	tcatgcatcc	4200
gttgatggac	acttgggttg	attccgtgtc	ctggctgctg	ggacagtgtc	gcgatgaaca	4260
cgagggtaca	gacgcctctc	ctacacgcta	atttcaactc	tttggatata	caccacagcag	4320
tgggattgct	ggatcagggtg	ggagctctat	ttccacattt	ttgagggaacc	tccctgccgt	4380
ctcccatggt	ggctgtgcca	acgacgttcc	cagggacaga	gtgcaacggg	cccctttcct	4440
ccatgtcctc	gccaacactc	gctatctttt	gcgttttgat	gacagtcac	ccaatagggtg	4500
ccagttggta	cctcctgtgg	tttttatttg	attttcctga	tgattagtga	tgctggacgt	4560
tatttcgtct	acacttcggc	cacttacatg	ttttccttcg	agacacgcag	attcagggtcc	4620
tttgcaacgtt	ttaaaaattt	ttttgtttgt	ttttgttatt	gagttgaatt	ccttctacaa	4680
tttgcaaatt	aactcctcat	catatacatg	gattgcaaatt	accccgccct	ccccctgggg	4740
ttttgccttt	tccctgcaaa	tactcccgcc	tcccatggg	ggttgccctt	tccctgccaa	4800
tacccccacc	tcccatggg	ggttgccctt	tccctgcaaa	tacccccacc	tccctgtggg	4860
ttctgccttt	tccctgcaaa	tacccccgcc	tccctgtggg	ggttgccctt	tactctgttt	4920

CCAGGAGCTG

ggtttccttt	gcggaagctt	tctggtttgt	tgcactctca	ctgtctattt	ttgcttctgt	4980
tgctctgtct	tgtggggcca	tatttttaaa	aaatcattgc	ccggaccagc	ctcaagaagt	5040
tttctccta	cgttttcttc	taagagtttt	atggtgtcgg	gtcttaggtt	tgaatcttta	5100
atccgtgttg	agttgatttt	cgtaggtggt	gtcggatgag	gccctttcat	cctcctccac	5160
ttttcccagc	accacctatt	gaggatgccc	ctttcccogt	cgtgtgtcct	tggcgccctt	5220
gctgaaggtc	agttggccgt	aactgtgcat	ggggaccctt	cctggccccc	ctggtgccc	5280
gtgccccata	tgtcccaccc	cctcccttac	tttttctcca	tggcatgaat	caccccagac	5340
ctactataca	aaatttatcc	tatttatttt	tatttattta	tttatttttg	agatggagtc	5400
tcactctgtc	acccaggctg	gagtgcagtg	gcacgatctc	ggctcactgc	aaagctccgcc	5460
tcccagggtc	acgccattct	cctgcctcag	cctcccaagt	agctgggatt	acgggcgcgc	5520
gccaccatgc	ccggctaagt	tttttgtttt	tttcgtagag	acagagtttc	cctatgtttg	5580
ccccagggtg	gtctcgaact	cctgggctca	agtaatcctt	ccacttcggc	ctcccaaagt	5640
gctgggatta	caggcatgag	ccattcgccc	cggcctattt	tttttttttc	agacagagtt	5700
tcactcttgt	cacccaggct	gaaatgcatt	gcaatgatct	tggctcactg	caacttccac	5760
ctcccagggt	caaaaggattt	ttctggcctc	agcctcccga	ggagctggga	ttacagtgtg	5820
caaccaccac	accgggctaa	aatttttggg	attttttttt	tttactagag	acagggttca	5880
acaatgctgg	tcaggctggt	ctcgaattcc	tgacctcaag	tgatcctccc	acctcggcct	5940
cccaaagtgc	tgggattaca	ggcgtgagcc	gccatgcctg	gccatggata	ttgtaaagtgt	6000
tcttgtttgt	tgtatgtttt	cctcactggg	ctgtgcactc	ctgaggcgcg	ggcatctgtc	6060
ccattcttca	gtgctgggtc	ccctgtgtct	gggacagtgt	atacatacag	cagggtgcata	6120
atcagtcttg	actggaaggg	tgaggggagt	aacgcacatg	gcagtcatgt	gactatgtgt	6180
ctgagaagca	taactcactt	aatcttgaag	ttcacttatg	gattgaagtg	tgcggttcag	6240
tgacttttaa	tatatattacc	gagtttgtta	accatcacca	ccatctaatt	ttaaatcatt	6300
ttcatcatcc	ctaaaagaaa	cttcagaccc	actagctgtc	cctcccccta	ttcctcccac	6360
cccagccctg	gtcctggccg	caggctgtct	acctgcactc	ctctgtggat	ctgccggttg	6420
tggacatttc	acacacctgc	gtgcagtcct	ctgtgcctgc	ctctttcact	cgtgtgtgat	6480
tttaagttca	cccatgttgt	catctatata	ggctacttact	tctttttttt	ttttgtagat	6540
gaagtcttgt	tcttgtcacc	caggctggag	tgacgtggcg	tgatctcgcc	tcacagcaac	6600
ttctgcctct	ggggttcaag	tgattctcct	gccttagcct	cccaagtagc	tgggactaca	6660
ggtttgcacc	accatgtcct	gctaattttt	ttttttttgt	atttttaata	gagacagggt	6720
ttctcctcat	tggccaggct	ggtctcgaac	tcttgacctc	agacgatcca	cctgcctcag	6780
cctcccgaag	tgttgggatt	acaggcacga	gccactgtgc	ccggccatca	ttccttttta	6840
ctgctgacta	atagtctgct	gtgtgaatcc	accgctagaa	accactcat	cagttgatgg	6900
tcatgtgggt	tgtttctgct	attcgcttat	tatgaacagt	gctggaataa	acgttcctgt	6960
gcactcttgg	gcatacgctt	aggagtggaa	ctgctgggtc	aaatggtgac	tttacgttta	7020
acgttctgag	gagccgccag	gcgttttaac	acagtgactg	caccatttca	cattcctgcc	7080
aacaatgtgt	gagaattcca	atttctctac	atccccaa	ttttccttta	aaaaaaagaa	7140
aaaagaaaca	tagccatcta	agtggatgtg	gagcagactg	tccctctggt	tgggttttgc	7200
gttgctttta	tggctcatga	tgtctgagtc	tctctccatg	tgtcatggg	gattcgtata	7260
tctacttttg	gaaatgctta	ttcaagtcct	ttgtccacat	ttgactgggt	tgtctgtctt	7320
tttatttcat	ttactacgat	gacagccctt	acatggaagg	attttgtttt	tgtaatccca	7380
ttaccccgag	tgtgaaatga	attgccagtt	gctcaaggcc	ttcagctctt	agggaggagc	7440
ctggacctgg	agctgctccg	ggctctggca	aagctccaat	ccggcctca	gtccttgagg	7500
cctggtcctc	acccagcttt	ctccttccac	cgtgccatgg	aggaagcccg	acctccctgc	7560
acggctggcc	tggggttgtt	cacgactgag	tccagggtgc	cccagaacgg	atgtcactgg	7620
tcacagtgtt	cctggtaata	ggtgacccca	ggcacagggt	gttccctgat	ataggttaacc	7680
caggcacagg	tgtcccagtc	acaggtgtct	ccaggcacag	gtgtccccag	tcacagggtgt	7740
cccaggtcac	aggcgtcccc	aggcacaggt	gtccctgggt	acagatgtcc	ccaggcacag	7800
gtgtcccagg	cacaggtgtc	tccaggcaca	ggcgtcccag	gtcacagggt	tccccgggtc	7860
cagggtgtccc	tgggtcacagg	tgtctccagg	cacaggtgtc	cctggtcaca	ggtgtccccg	7920
gtcacagggt	tcccagggtca	cagggtgtccc	caggcacagg	tgtccccggt	cacaggtgtc	7980
tccggtcaca	ggtgtcccca	ggcatagggt	tccctgggtc	caggcaccca	tgggtcacagg	8040
tgtccccagg	cacaggtgtc	ctggtccacag	gtgtcccagg	cacagctgtc	cccgttcaca	8100
ggtgtctcca	ggcacagggt	tccccggtca	caggtgtccc	caggcacagg	tgtccccgtc	8160
acaggtgtcc	ccaggcacag	gagttcctgg	tcacagggtgt	ccccaggcac	aggcagccac	8220
aggaagccga	tgcattggaac	agagagaaac	agagacacaa	agaaaagaga	gtgagagaca	8280
gaagaaatgg	gaaacagaaa	tgggttgaga	aaagcatcca	gtagacatga	atagagagga	8340

```
<210> 2
<211> 1581
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> CDS
<222> (3)...(1310)
```

```

<400> 2
tc acg agc tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga      47
   Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg
       1               5               10               15

```

gct gcc cac gtc gtc tcc agg aag gga ccc ggg tcc acg agc tgc cca 95
Ala Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro
20 25 30

cgt cct ctc cag gaa agg acc cgg gtc cac gag ctg gcc acg tcc tct 143
Arg Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser
35 40 45

gca Ala	gga Gly	agg Arg 50	gac Asp	ccc Pro	ggg Gly	tcc Ser	acg Thr 55	agc Ser	tgc Cys	cca Pro	cgt Arg	cct Pro 60	ctc Leu	cag Gln	gaa Glu	191
ggg Gly	acc Thr 65	ccg Pro	ggt Gly	tca Ser	cga Arg	gct Ala 70	gcc Ala	cac His	gtc Val	ctc Leu	tcc Ser 75	agg Arg	aag Lys	gga Gly	ccc Pro	239
cgg Arg 80	gtc Val	cac His	gag Glu	ctg Leu	ccc Pro 85	acg Thr	tcc Ser	tct Ser	cca Pro	gga Gly 90	agg Arg	gac Asp	ccc Pro	ggg Gly	tcc Ser 95	287
acg Thr	aac Asn	tgc Cys	cca Pro	cgt Arg 100	cct Pro	ctc Leu	cag Gln	gaa Glu	ggg Gly 105	acc Thr	ccg Pro	ggt Gly	tca Ser	cga Arg 110	gct Ala	335
gcc Ala	cac His	gtc Val 115	ctc Leu	tcc Ser	agg Arg	agg Arg	gga Gly 120	cac His	cgg Arg	gtt Val	cac His	gag Glu	ctg Leu 125	ccc Pro	acg Thr	383
ccc Pro	tct Ser	cca Pro 130	gga Gly	agg Arg	gac Asp	ccc Pro	ggg Gly 135	ttc Phe	atg Met	agc Ser	tgc Cys	cca Pro 140	cgt Arg	cct Pro	ctc Leu	431
cag Gln 145	gaa Glu	ggg Gly	acc Thr	cgg Arg	gtc Val	cac His 150	gaa Glu	ctg Leu	ccc Pro	acg Thr 155	ccc Pro	tct Ser	cca Pro	gga Gly	ggg Gly	479
gac Asp 160	ccg Pro	ggt Gly	cca Pro	cga Arg	gct Ala 165	gcc Ala	cac His	gtc Val	gtc Val	aac Asn 170	ggg Gly	aag Lys	gga Gly	ccc Pro	ggg Gly 175	527
tcc Ser	acg Thr	agc Ser	tgc Cys	cca Pro 180	cgt Arg	cct Pro	ctc Leu	cag Gln	gaa Glu 185	ggg Gly	acc Thr	cgg Arg	gtc Val	cac His 190	gaa Glu	575
ctg Leu	ccc Pro	acg Thr	cgc Arg 195	tct Ser	cca Pro	gga Gly	ggg Gly	gac Asp 200	acc Thr	ggg Gly	ttc Phe	acg Thr 205	agc Ser	tgc Cys	cca Pro	623
cgc Arg	cct Pro	ctc Leu 210	cag Gln	gaa Glu	ggg Gly	acc Thr	ccg Pro 215	ggt Gly	tca Ser	cga Arg	gct Ala	gcc Ala 220	cac His	gtc Val	ctc Leu	671
tcc Ser	agg Arg 225	agg Arg	gga Gly	cac His	cgg Arg	gtt Val 230	cac His	gag Glu	ctg Leu	ccc Pro	acg Thr 235	tcc Ser	tct Ser	cca Pro	gga Gly	719
ggg Gly 240	gac Asp	acc Thr	ggg Gly	ttc Phe	acg Thr 245	agc Ser	tgc Cys	cca Pro	cgc Arg	cct Pro 250	ctc Leu	cag Gln	gag Glu	ggg Gly	aca Thr 255	767
ccg Pro	ggt Gly	tca Ser	cga Arg	gct Ala 260	gcc Ala	cac His	gtc Val	ctc Leu	tcc Ser 265	agg Arg	aag Lys	gga Gly	ccc Pro	ggg Gly	tcc Ser 270	815

acg agc tgc cca cgt cct ctc cag gag ggg aca ccg ggt tca cga gct 863
 Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala
 275 280 285

gcc cac gca ctt tcc agg aag gga ccc cgg gtt cag gtc tcc tgc cgg 911
 Ala His Ala Leu Ser Arg Lys Gly Pro Arg Val Gln Val Ser Cys Arg
 290 295 300

ccc aca tcg tgc ctt tgt gta aat cag aag aaa gat gag gaa cag gcc 959
 Pro Thr Ser Cys Leu Cys Val Asn Gln Lys Lys Asp Glu Glu Gln Ala
 305 310 315

ctc ctc tct ctc cag gca ggc ttt ggt gga ggg gct gga tct cct gcc 1007
 Leu Leu Ser Leu Gln Ala Gly Phe Gly Gly Ala Gly Ser Pro Ala
 320 325 330 335

gca cct tcc ctg gca ggg cac cct gtg ctt gag ccc cag aac tgc agg 1055
 Ala Pro Ser Leu Ala Gly His Pro Val Leu Glu Pro Gln Asn Cys Arg
 340 345 350

cgg ccg gca gag aag ggg tcc atg atg gcg cct cgg tgc gca gcc ttg 1103
 Arg Pro Ala Glu Lys Gly Ser Met Met Ala Pro Arg Cys Ala Ala Leu
 355 360 365

gac ctg ccc cca tgg acc tgg gaa cct ccc ggc tct tcc cac tcg gga 1151
 Asp Leu Pro Pro Trp Thr Trp Glu Pro Pro Gly Ser Ser His Ser Gly
 370 375 380

aag gaa ggc tct ggg cat gga ggt cgg cca ggc ccc atc ccc gta ccc 1199
 Lys Glu Gly Ser Gly His Gly Gly Arg Pro Gly Pro Ile Pro Val Pro
 385 390 395

tgg ccc ttc ttc ctg ctt cct gtt tgt cac tgc ccc ggg gcc ttt gca 1247
 Trp Pro Phe Phe Leu Leu Pro Val Cys His Cys Pro Gly Ala Phe Ala
 400 405 410 415

cct gca ttc cct ctc tct aga cag ggt ttc tcc tca ttg gcc agg ctg 1295
 Pro Ala Phe Pro Leu Ser Arg Gln Gly Phe Ser Ser Leu Ala Arg Leu
 420 425 430

gtc tcg aac tcc tga cctcagacga tccacctgcc tcagcctccc gaagtgttgg 1350
 Val Ser Asn Ser *
 435

gattacaggc acgagccact gtgcccggcc atcattcctt tttactgctg actaatagtc 1410
 tgctgtgtga atccaccgct agaaaccac tcatcagttg atggatcatgt gggttgcttc 1470
 tgctattcgc ttattatgaa cagtgtctgga ataaacgttc ctgtgcactc ttgggcatac 1530
 gcctaggagt ggaactgctg ggtcaaaaaa aaaaaaaaaa aaaaaaaaaa a 1581

<210> 3
 <211> 435
 <212> PRT
 <213> Homo sapiens

Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly	Ser	Arg	Ala
1				5					10					15	
Ala	His	Val	Val	Ser	Arg	Lys	Gly	Pro	Gly	Ser	Thr	Ser	Cys	Pro	Arg
			20					25					30		
Pro	Leu	Gln	Glu	Arg	Thr	Arg	Val	His	Glu	Leu	Ala	Thr	Ser	Ser	Ala
		35					40					45			
Gly	Arg	Asp	Pro	Gly	Ser	Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly
	50					55					60				
Thr	Pro	Gly	Ser	Arg	Ala	Ala	His	Val	Leu	Ser	Arg	Lys	Gly	Pro	Arg
65					70					75					80
Val	His	Glu	Leu	Pro	Thr	Ser	Ser	Pro	Gly	Arg	Asp	Pro	Gly	Ser	Thr
				85					90					95	
Asn	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly	Ser	Arg	Ala	Ala
			100					105					110		
His	Val	Leu	Ser	Arg	Arg	Gly	His	Arg	Val	His	Glu	Leu	Pro	Thr	Pro
		115					120					125			
Ser	Pro	Gly	Arg	Asp	Pro	Gly	Phe	Met	Ser	Cys	Pro	Arg	Pro	Leu	Gln
	130					135					140				
Glu	Gly	Thr	Arg	Val	His	Glu	Leu	Pro	Thr	Pro	Ser	Pro	Gly	Gly	Asp
145					150					155					160
Pro	Gly	Pro	Arg	Ala	Ala	His	Val	Val	Asn	Gly	Lys	Gly	Pro	Gly	Ser
				165					170					175	
Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Arg	Val	His	Glu	Leu
			180					185					190		
Pro	Thr	Arg	Ser	Pro	Gly	Gly	Asp	Thr	Gly	Phe	Thr	Ser	Cys	Pro	Arg
		195					200					205			
Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly	Ser	Arg	Ala	Ala	His	Val	Leu	Ser
	210					215					220				
Arg	Arg	Gly	His	Arg	Val	His	Glu	Leu	Pro	Thr	Ser	Ser	Pro	Gly	Gly
225					230					235					240
Asp	Thr	Gly	Phe	Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Pro
				245					250					255	
Gly	Ser	Arg	Ala	Ala	His	Val	Leu	Ser	Arg	Lys	Gly	Pro	Gly	Ser	Thr
			260					265					270		
Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly	Ser	Arg	Ala	Ala
		275					280					285			
His	Ala	Leu	Ser	Arg	Lys	Gly	Pro	Arg	Val	Gln	Val	Ser	Cys	Arg	Pro
	290					295					300				
Thr	Ser	Cys	Leu	Cys	Val	Asn	Gln	Lys	Lys	Asp	Glu	Gln	Ala	Leu	
305					310					315					320
Leu	Ser	Leu	Gln	Ala	Gly	Phe	Gly	Gly	Gly	Ala	Gly	Ser	Pro	Ala	Ala
			325						330					335	
Pro	Ser	Leu	Ala	Gly	His	Pro	Val	Leu	Glu	Pro	Gln	Asn	Cys	Arg	Arg
			340					345					350		
Pro	Ala	Glu	Lys	Gly											

<210> 4
 <211> 1441
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (3)...(1169)

<400> 4
 tc acg agc tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga 47
 Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg
 1 5 10 15

gct gcc cac gtc gtc tcc agg aag gga ccc ggg tcc acg agc tgc cca 95
 Ala Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro
 20 25 30

cgt cct ctc cag gaa agg acc cgg gtc cac gag ctg gcc acg tcc tct 143
 Arg Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser
 35 40 45

gca gga agg gac ccc ggg tcc acg agc tgc cca cgt cct ctc cag gaa 191
 Ala Gly Arg Asp Pro Gly Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu
 50 55 60

ggg acc ccg ggt tca cga gct gcc cac gtc ctc tcc agg aag gga ccc 239
 Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro
 65 70 75

cgg gtc cac gag ctg ccc acg tcc tct cca gga agg gac ccc ggg tcc 287
 Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly Arg Asp Pro Gly Ser
 80 85 90 95

acg aac tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga gct 335
 Thr Asn Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala
 100 105 110

gcc cac gtc ctc tcc agg agg gga cac cgg gtt cac gag ctg ccc acg 383
 Ala His Val Leu Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr
 115 120 125

ccc tct cca gga agg gac ccc ggg ttc atg agc tgc cca cgt cct ctc 431
 Pro Ser Pro Gly Arg Asp Pro Gly Phe Met Ser Cys Pro Arg Pro Leu
 130 135 140

cag gaa ggg acc cgg gtc cac gaa ctg ccc acg ccc tct cca gga ggg 479
 Gln Glu Gly Thr Arg Val His Glu Leu Pro Thr Pro Ser Pro Gly Gly
 145 150 155

gac ccg ggt cca cga gct gcc cac gtc gtc aac ggg aag gga ccc ggg 527
 Asp Pro Gly Pro Arg Ala Ala His Val Val Asn Gly Lys Gly Pro Gly
 160 165 170 175

CCDS:CCDS1000.1

tcc	acg	agc	tgc	cca	cgt	cct	ctc	cag	gaa	ggg	acc	cgg	gtc	cac	gaa		575
Ser	Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Arg	Val	His	Glu		
				180					185					190			
ctg	ccc	acg	cgc	tct	cca	gga	ggg	gac	acc	ggg	ttc	acg	agc	tgc	cca		623
Leu	Pro	Thr	Arg	Ser	Pro	Gly	Gly	Asp	Thr	Gly	Phe	Thr	Ser	Cys	Pro		
			195					200					205				
cgc	cct	ctc	cag	gaa	ggg	acc	ccg	ggt	tca	cga	gct	gcc	cac	gtc	ctc		671
Arg	Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly	Ser	Arg	Ala	Ala	His	Val	Leu		
		210					215					220					
tcc	agg	agg	gga	cac	cgg	gtt	cac	gag	ctg	ccc	acg	tcc	tct	cca	gga		719
Ser	Arg	Arg	Gly	His	Arg	Val	His	Glu	Leu	Pro	Thr	Ser	Ser	Pro	Gly		
	225					230					235						
ggg	gac	acc	ggg	ttc	acg	agc	tgc	cca	cgc	cct	ctc	cag	gag	ggg	aca		767
Gly	Asp	Thr	Gly	Phe	Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr		
240					245					250					255		
ccg	ggt	tca	cga	gct	gcc	cac	gtc	ctc	tcc	agg	aag	gga	ccc	ggg	tcc		815
Pro	Gly	Ser	Arg	Ala	Ala	His	Val	Leu	Ser	Arg	Lys	Gly	Pro	Gly	Ser		
				260					265					270			
acg	agc	tgc	cca	cgt	cct	ctc	cag	gag	ggg	aca	ccg	ggt	tca	cga	gct		863
Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly	Ser	Arg	Ala		
			275					280					285				
gcc	cac	gca	ctt	tcc	agg	aag	gga	ccc	cgg	gtt	cag	gtc	tcc	tgc	cgg		911
Ala	His	Ala	Leu	Ser	Arg	Lys	Gly	Pro	Arg	Val	Gln	Val	Ser	Cys	Arg		
		290					295					300					
ccc	aca	tcg	tgc	ctt	tgt	gta	aat	cag	aag	aaa	gat	gag	gaa	cag	gcc		959
Pro	Thr	Ser	Cys	Leu	Cys	Val	Asn	Gln	Lys	Lys	Asp	Glu	Glu	Gln	Ala		
	305					310					315						
ctc	ctc	tct	ctc	cag	gca	ggc	ttt	ggt	gga	ggg	gct	gga	tct	cct	gcc		1007
Leu	Leu	Ser	Leu	Gln	Ala	Gly	Phe	Gly	Gly	Gly	Ala	Gly	Ser	Pro	Ala		
320					325					330					335		
gca	cct	tcc	ctg	gca	ggg	cac	cct	gtg	ctt	gag	ccc	cag	aac	tgc	agg		1055
Ala	Pro	Ser	Leu	Ala	Gly	His	Pro	Val	Leu	Glu	Pro	Gln	Asn	Cys	Arg		
				340					345					350			
cgg	ccg	gca	gag	aag	ggg	tcc	atg	atg	gcg	cct	cgg	tgc	gca	gcc	ttg		1103
Arg	Pro	Ala	Glu	Lys	Gly	Ser	Met	Met	Ala	Pro	Arg	Cys	Ala	Ala	Leu		
			355					360					365				
gac	ctg	ccc	cca	tgg	acc	tgg	aga	cag	ggt	ttc	tcc	tca	ttg	gcc	agg		1151
Asp	Leu	Pro	Pro	Trp	Thr	Trp	Arg	Gln	Gly	Phe	Ser	Ser	Leu	Ala			

actaatagtc tgctgtgtga atccaccgct agaaacccac tcatacagttg atgggtcatgt 1319
 ggggttgcttc tgctattcgc ttattatgaa cagtgtctgga ataaacgttc ctgtgcactc 1379
 ttgggcatac gcctaggagt ggaactgctg ggtcaaaaaa aaaaaaaaaa aaaaaaaaaa 1439
 aa 1441

<210> 5
 <211> 388
 <212> PRT
 <213> Homo sapiens

<400> 5
 Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala
 1 5 10 15
 Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro Arg
 20 25 30
 Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser Ala
 35 40 45
 Gly Arg Asp Pro Gly Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly
 50 55 60
 Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro Arg
 65 70 75 80
 Val His Glu Leu Pro Thr Ser Ser Pro Gly Arg Asp Pro Gly Ser Thr
 85 90 95
 Asn Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala
 100 105 110
 His Val Leu Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr Pro
 115 120 125
 Ser Pro Gly Arg Asp Pro Gly Phe Met Ser Cys Pro Arg Pro Leu Gln
 130 135 140
 Glu Gly Thr Arg Val His Glu Leu Pro Thr Pro Ser Pro Gly Gly Asp
 145 150 155 160
 Pro Gly Pro Arg Ala Ala His Val Val Asn Gly Lys Gly Pro Gly Ser
 165 170 175
 Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Arg Val His Glu Leu
 180 185 190
 Pro Thr Arg Ser Pro Gly Gly Asp Thr Gly Phe Thr Ser Cys Pro Arg
 195 200 205
 Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser
 210 215 220
 Arg Arg Gly His Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly Gly
 225 230 235 240
 Asp Thr Gly Phe Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro
 245 250 255
 Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro Gly Ser Thr
 260 265 270
 Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala
 275 280 285
 His Ala Leu Ser Arg Lys Gly Pro Arg Val Gln Val Ser Cys Arg Pro
 290 295 300
 Thr Ser Cys Leu Cys Val Asn Gln Lys Lys Asp Glu Glu Gln Ala Leu
 305 310 315 320
 Leu Ser Leu Gln Ala Gly Phe Gly Gly Gly Ala Gly Ser Pro Ala Ala
 325 330 335
 Pro Ser Leu Ala Gly His Pro Val Leu Glu Pro Gln Asn Cys Arg Arg
 340 345 350
 Pro Ala Glu Lys Gly Ser Met Met Ala Pro Arg Cys Ala Ala Leu Asp
 355 360 365

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Leu Pro Pro Trp Thr Trp Arg Gln Gly Phe Ser Ser Leu Ala Arg Leu
 370 375 380
 Val Ser Asn Ser
 385

<210> 6
 <211> 1576
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (3)...(1193)

<400> 6
 tc acg agc tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga 47
 Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg
 1 5 10 15
 gct gcc cac gtc gtc tcc agg aag gga ccc ggg tcc acg agc tgc cca 95
 Ala Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro
 20 25 30
 cgt cct ctc cag gaa agg acc cgg gtc cac gag ctg gcc acg tcc tct 143
 Arg Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser
 35 40 45
 gca gga agg gac ccc ggg tcc acg agc tgc cca cgt cct ctc cag gaa 191
 Ala Gly Arg Asp Pro Gly Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu
 50 55 60
 ggg acc ccg ggt tca cga gct gcc cac gtc ctc tcc agg aag gga ccc 239
 Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro
 65 70 75
 cgg gtc cac gag ctg ccc acg tcc tct cca gga agg gac ccc ggg tcc 287
 Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly Arg Asp Pro Gly Ser
 80 85 90 95
 acg aac tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga gct 335
 Thr Asn Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala
 100 105 110
 gcc cac gtc ctc tcc agg agg gga cac cgg gtt cac gag ctg ccc acg 383
 Ala His Val Leu Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr
 115 120 125
 ccc tct cca gga agg gac ccc ggg ttc atg agc tgc cca cgt cct ctc 431
 Pro Ser Pro Gly Arg Asp Pro Gly Phe Met Ser Cys Pro Arg Pro Leu
 130 135 140
 cag gaa ggg acc cgg gtc cac gaa ctg ccc acg ccc tct cca gga ggg 479
 Gln Glu Gly Thr Arg Val His Glu Leu Pro Thr Pro Ser Pro Gly Gly
 145 150 155

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted November 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gac Asp 160	ccg Pro	ggt Gly	cca Pro	cga Arg	gct Ala 165	gcc Ala	cac His	gtc Val	gtc Val	aac Asn 170	ggg Gly	aag Lys	gga Gly	ccc Pro	ggg Gly 175	527
tcc Ser	acg Thr	agc Ser	tgc Cys	cca Pro 180	cg Arg	cct Pro	ctc Leu	cag Gln	gaa Glu 185	ggg Gly	acc Thr	cgg Arg	gtc Val	cac His 190	gaa Glu	575
ctg Leu	ccc Pro	acg Thr	cgc Arg 195	tct Ser	cca Pro	gga Gly	ggg Gly	gac Asp 200	acc Thr	ggg Gly	ttc Phe	acg Thr	agc Ser 205	tgc Cys	cca Pro	623
cgc Arg	cct Pro	ctc Leu 210	cag Gln	gaa Glu	ggg Gly	acc Thr	ccg Pro	ggt Gly 215	tca Ser	cga Arg	gct Ala	gcc Ala 220	cac His	gtc Val	ctc Leu	671
tcc Ser 225	agg Arg	agg Arg	gga Gly	cac His	cgg Arg	gtt Val 230	cac His	gag Glu	ctg Leu	ccc Pro	acg Thr 235	tcc Ser	tct Ser	cca Pro	gga Gly	719
ggg Gly 240	gac Asp	acc Thr	ggg Gly	ttc Phe	acg Thr 245	agc Ser	tgc Cys	cca Pro	cgc Arg	cct Pro 250	ctc Leu	cag Gln	gag Glu	ggg Gly 255	aca Thr	767
ccg Pro	ggt Gly	tca Ser	cga Arg 260	gct Ala	gcc Ala	cac His	gtc Val	ctc Leu	tcc Ser 265	agg Arg	aag Lys	gga Gly	ccc Pro	ggg Gly 270	tcc Ser	815
acg Thr	agc Ser	tgc Cys	cca Pro 275	cg Arg	cct Pro	ctc Leu	cag Gln	gag Glu 280	ggg Gly	aca Thr	ccg Pro	ggt Gly	tca Ser 285	cga Arg	gct Ala	863
gcc Ala	cac His	gca Ala 290	ctt Leu	tcc Ser	agg Arg	aag Lys	gga Gly 295	ccc Pro	cgg Arg	gtt Val	cag Gln	gtc Val 300	tcc Ser	tgc Cys	cgg Arg	911
ccc Pro 305	aca Thr	tcg Ser	tgc Cys	ctt Leu	tgt Cys	gta Val 310	aat Asn	cag Gln	aag Lys	aaa Lys 315	gat Asp	gag Glu	gaa Glu	cag Gln	gcc Ala	959
ctc Leu 320	ctc Leu	tct Ser	ctc Leu	cag Gln	gca Ala 325	ggc Gly	ttt Phe	ggt Gly	gga Gly	ggg Gly 330	gct Ala	gga Gly	tct Ser	cct Pro	gcc Ala 335	1007
gca Ala	cct Pro	tcc Ser	ctg Leu 340	gca Ala	ggg Gly	cac His	cct Pro	gtg Val	ctt Leu 345	gag Glu	ccc Pro	cag Gln	aac Asn	tgc Cys 350	agg Arg	1055
cgg Arg	ccg Pro	gca Ala	gag Glu 355	aag Lys	ggg Gly	tcc Ser	atg Met	atg Met 360	gcg Ala	cct Pro	cgg Arg	tgc Cys	gca Ala	gcc Ala	ttg Leu	1103
gac Asp	ctg Leu	ccc Pro 370	cca Pro	tgg Trp	acc Thr	tgg Trp	gaa Glu 375	cct Pro	ccc Pro	ggc Gly	tct Ser	tcc Ser 380	cac His	tcg Ser	gga Gly	1151

1193

caattcagga	cgggtgtgcac	tcaagggtatg	cagcatcaca	acctgacaca	cgtaggcatt	1253
gtgaaatgag	tcccacaatt	gggctaatta	acacacccat	caccttacat	ggttacttct	1313
ttctgtggtg	agaacactaa	attttaaata	gaggacacac	agcctgggca	acatagttag	1373
acctgtctc	tacaaatata	aaaaaattat	ctggacgtgg	tgggtgcacac	ctgtgggtccc	1433
agctacttgg	gaagctgagg	ctggagaatc	acttgagcct	gggaggcgga	ggttgctggtg	1493
cactccagcc	tgggcgacag	agggaggccc	tatctcaaaa	taaataaata	aaggacacat	1553
tcttatcaaa	aaaaaaaaaa	aaa				1576

```
<210> 7
<211> 396
<212> PRT
<213> Homo sapiens
```

<400> 7																
Thr 1	Ser	Cys	Pro	Arg 5	Pro	Leu	Gln	Glu	Gly 10	Thr	Pro	Gly	Ser	Arg 15	Ala	
Ala	His	Val	Val	Ser	Arg	Lys	Gly	Pro	Gly 25	Ser	Thr	Ser	Cys 30	Pro	Arg	
Pro	Leu	Gln	Glu	Arg	Thr	Arg	Val	His	Glu 40	Leu	Ala	Thr 45	Ser	Ser	Ala	
Gly	Arg	Asp	Pro	Gly	Ser	Thr 55	Ser	Cys	Pro	Arg	Pro 60	Leu	Gln	Glu	Gly	
Thr 65	Pro	Gly	Ser	Arg 70	Ala	His	Val	Leu	Ser 75	Arg	Lys	Gly	Pro	Arg 80		
Val	His	Glu	Leu	Pro 85	Thr	Ser	Ser	Pro	Gly 90	Arg	Asp	Pro	Gly	Ser 95	Thr	
Asn	Cys	Pro	Arg	Pro 100	Leu	Gln	Glu	Gly 105	Thr	Pro	Gly	Ser	Arg 110	Ala	Ala	
His	Val	Leu	Ser	Arg	Arg	Gly	His 120	Arg	Val	His	Glu	Leu 125	Pro	Thr	Pro	
Ser	Pro	Gly	Arg	Asp	Pro	Gly 135	Phe	Met	Ser	Cys	Pro 140	Arg	Pro	Leu	Gln	
Glu 145	Gly	Thr	Arg	Val	His 150	Glu	Leu	Pro	Thr	Pro 155	Ser	Pro	Gly	Gly	Asp 160	
Pro	Gly	Pro	Arg	Ala 165	His	Val	Val	Val	Asn 170	Gly	Lys	Gly	Pro	Gly 175	Ser	
Thr	Ser	Cys	Pro	Arg 180	Pro	Leu	Gln	Glu 185	Gly	Thr	Arg	Val	His 190	Glu	Leu	
Pro	Thr	Arg	Ser	Pro	Gly	Gly	Asp 200	Thr	Gly	Phe	Thr	Ser 205	Cys	Pro	Arg	
Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly 215	Ser	Arg	Ala	Ala 220	His	Val	Leu	Ser	
Arg 225	Arg	Gly	His	Arg	Val	His	Glu	Leu	Pro	Thr 235	Ser	Ser	Pro	Gly	Gly 240	
Asp	Thr	Gly	Phe	Thr 245	Ser	Cys	Pro	Arg	Pro 250	Leu	Gln	Glu	Gly	Thr 255	Pro	
Gly	Ser	Arg	Ala	Ala 260	His	Val	Leu	Ser 265	Arg	Lys	Gly	Pro	Gly 270	Ser	Thr	
Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu 280	Gly	Thr	Pro	Gly	Ser 285	Arg	Ala	Ala	
His	Ala	Leu	Ser	Arg	Lys	Gly 295	Pro	Arg	Val	Gln	Val 300	Ser	Cys	Arg	Pro	

[illegible]

Thr Ser Cys Leu Cys Val Asn Gln Lys Lys Asp Glu Glu Gln Ala Leu
 305 310 315 320
 Leu Ser Leu Gln Ala Gly Phe Gly Gly Gly Ala Gly Ser Pro Ala Ala
 325 330 335
 Pro Ser Leu Ala Gly His Pro Val Leu Glu Pro Gln Asn Cys Arg Arg
 340 345 350
 Pro Ala Glu Lys Gly Ser Met Met Ala Pro Arg Cys Ala Ala Leu Asp
 355 360 365
 Leu Pro Pro Trp Thr Trp Glu Pro Pro Gly Ser Ser His Ser Gly Lys
 370 375 380
 Glu Gly Ser Gly His Gly Ala Leu Leu Arg Tyr Ser
 385 390 395

<210> 8

<211> 2010

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (3)...(1247)

<400> 8

tc acg agc tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga	47
Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg	
1 5 10 15	
gct gcc cac gtc gtc tcc agg aag gga ccc ggg tcc acg agc tgc cca	95
Ala Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro	
20 25 30	
cgt cct ctc cag gaa agg acc cgg gtc cac gag ctg gcc acg tcc tct	143
Arg Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser	
35 40 45	
gca gga agg gac ccc ggg tcc acg agc tgc cca cgt cct ctc cag gaa	191
Ala Gly Arg Asp Pro Gly Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu	
50 55 60	
ggg acc ccg ggt tca cga gct gcc cac gtc ctc tcc agg aag gga ccc	239
Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro	
65 70 75	
cgg gtc cac gag ctg ccc acg tcc tct cca gga agg gac ccc ggg tcc	287
Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly Arg Asp Pro Gly Ser	
80 85 90 95	
acg aac tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga gct	335
Thr Asn Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala	
100 105 110	
gcc cac gtc ctc tcc agg agg gga cac cgg gtt cac gag ctg ccc acg	383
Ala His Val Leu Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr	
115 120 125	

CCDS: M36123.1

431
479
527
575
623
671
719
767
815
863
911
959
1007
1055

cgg ccg gca gag aag ggg tcc atg atg gcg cct cgg tgc gca gcc ttg 1103
Arg Pro Ala Glu Lys Gly Ser Met Met Ala Pro Arg Cys Ala Ala Leu
355 360 365

gac ctg ccc cca tgg acc tgg atg cca gtg atg cct gag gtc tgc agg 1151
Asp Leu Pro Pro Trp Thr Trp Met Pro Val Met Pro Glu Val Cys Arg
370 375 380

gca gtg cat acg ctc acc gcc tgg ccg ctc agg agc ctg tgc ttg acc 1199
Ala Val His Thr Leu Thr Ala Trp Pro Leu Arg Ser Leu Cys Leu Thr
385 390 395

ccc aaa tcc gcc ccc caa ctc cct gtt acc ggc tca ctc ctt cca tga 1247
Pro Lys Ser Ala Pro Gln Leu Pro Val Thr Gly Ser Leu Leu Pro *
400 405 410

ggggccttcc ccagggacag ccgatgctct cctgatggct cctgcccttg cagagtgtg 1307
ccccgcctg cccacctggc ctggaccctc gctgagccc cctcagggct ctgcgcacc 1367
tcaaccacag cgttgttcc gcaggaacct ccggctctt cccactcggg aaaggaaggc 1427
tctgggcatg gaggtcgcc agggcccatc ccgtacctt ggcccttctt cctgcttctt 1487
gtttgtcact gccccggggc ctttgacact gcattccctc tctctgtgag tgtcctgggg 1547
cccgttacct acgtcaccgt cccaggatac cttttctttt ctttctctct ctccagcttt 1607
attgaggtat agttgacaat tcaggacggt gtgcaactca ggtatgcagc atcacaacct 1667
gacacacgta ggcattgtga aatgagctcc acaattgggc taattaacac acccatcacc 1727
ttacatgggt acttctttct gtggtgagaa cactaaattt taaatagagg acacacagcc 1787
tgggcaacat agtgagaccc tgtctctaca aatataaaaa aattatctgg acgtgggtggt 1847
gcacacctgt ggtcccagct acttggaag ctgaggctgg agaatactt gagcctggga 1907
ggcggagggt gcggtgcact ccagcctggg cgacagaggg aggccctatc tcaaaataaa 1967
taaataaagg acacattctt atcaaaaaaa aaaaaaaaaa aaa 2010

<210> 9
<211> 414
<212> PRT
<213> Homo sapiens

<400> 9
Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala
1 5 10 15
Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro Arg
20 25 30
Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser Ala
35 40 45
Gly Arg Asp Pro Gly Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly
50 55 60
Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro Arg
65 70 75 80
Val His Glu Leu Pro Thr Ser Ser Pro Gly Arg Asp Pro Gly Ser Thr
85 90 95
Asn Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala
100 105 110
His Val Leu Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr Pro
115 120 125
Ser Pro Gly Arg Asp Pro Gly Phe Met Ser Cys Pro Arg Pro Leu Gln
130 135 140
Glu Gly Thr Arg Val His Glu Leu Pro Thr Pro Ser Pro Gly Gly Asp
145 150 155 160

CCAGGGACAG

Pro Gly Pro Arg Ala Ala His Val Val Asn Gly Lys Gly Pro Gly Ser
 165 170 175
 Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Arg Val His Glu Leu
 180 185 190
 Pro Thr Arg Ser Pro Gly Gly Asp Thr Gly Phe Thr Ser Cys Pro Arg
 195 200 205
 Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser
 210 215 220
 Arg Arg Gly His Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly Gly
 225 230 235 240
 Asp Thr Gly Phe Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro
 245 250 255
 Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro Gly Ser Thr
 260 265 270
 Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala
 275 280 285
 His Ala Leu Ser Arg Lys Gly Pro Arg Val Gln Val Ser Cys Arg Pro
 290 295 300
 Thr Ser Cys Leu Cys Val Asn Gln Lys Lys Asp Glu Glu Gln Ala Leu
 305 310 315 320
 Leu Ser Leu Gln Ala Gly Phe Gly Gly Gly Ala Gly Ser Pro Ala Ala
 325 330 335
 Pro Ser Leu Ala Gly His Pro Val Leu Glu Pro Gln Asn Cys Arg Arg
 340 345 350
 Pro Ala Glu Lys Gly Ser Met Met Ala Pro Arg Cys Ala Ala Leu Asp
 355 360 365
 Leu Pro Pro Trp Thr Trp Met Pro Val Met Pro Glu Val Cys Arg Ala
 370 375 380
 Val His Thr Leu Thr Ala Trp Pro Leu Arg Ser Leu Cys Leu Thr Pro
 385 390 395 400
 Lys Ser Ala Pro Gln Leu Pro Val Thr Gly Ser Leu Leu Pro
 405 410

<210> 10
 <211> 1744
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (3)...(1352)

<400> 10
 tc acg agc tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga
 Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg
 1 5 10 15

gct gcc cac gtc gtc tcc agg aag gga ccc ggg tcc acg agc tgc cca
 Ala Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro
 20 25 30

cgt cct ctc cag gaa agg acc cgg gtc cac gag ctg gcc acg tcc tct
 Arg Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser
 35 40 45

47

95

143

gca gga agg gac ccc ggg tcc acg agc tgc cca cgt cct ctc cag gaa	191
Ala Gly Arg Asp Pro Gly Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu	
50 55 60	
ggg acc ccg ggt tca cga gct gcc cac gtc ctc tcc agg aag gga ccc	239
Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro	
65 70 75	
cgg gtc cac gag ctg ccc acg tcc tct cca gga agg gac ccc ggg tcc	287
Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly Arg Asp Pro Gly Ser	
80 85 90 95	
acg aac tgc cca cgt cct ctc cag gaa ggg acc ccg ggt tca cga gct	335
Thr Asn Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala	
100 105 110	
gcc cac gtc ctc tcc agg agg gga cac cgg gtt cac gag ctg ccc acg	383
Ala His Val Leu Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr	
115 120 125	
ccc tct cca gga agg gac ccc ggg ttc atg agc tgc cca cgt cct ctc	431
Pro Ser Pro Gly Arg Asp Pro Gly Phe Met Ser Cys Pro Arg Pro Leu	
130 135 140	
cag gaa ggg acc ccg gtc cac gaa ctg ccc acg ccc tct cca gga ggg	479
Gln Glu Gly Thr Arg Val His Glu Leu Pro Thr Pro Ser Pro Gly Gly	
145 150 155	
gac ccg ggt cca cga gct gcc cac gtc gtc aac ggg aag gga ccc ggg	527
Asp Pro Gly Pro Arg Ala Ala His Val Val Asn Gly Lys Gly Pro Gly	
160 165 170 175	
tcc acg agc tgc cca cgt cct ctc cag gaa ggg acc ccg gtc cac gaa	575
Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Arg Val His Glu	
180 185 190	
ctg ccc acg cgc tct cca gga ggg gac acc ggg ttc acg agc tgc cca	623
Leu Pro Thr Arg Ser Pro Gly Gly Asp Thr Gly Phe Thr Ser Cys Pro	
195 200 205	
cgc cct ctc cag gaa ggg acc ccg ggt tca cga gct gcc cac gtc ctc	671
Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu	
210 215 220	
tcc agg agg gga cac ccg gtt cac gag ctg ccc acg tcc tct cca gga	719
Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly	
225 230 235	
ggg gac acc ggg ttc acg agc tgc cca cgc cct ctc cag gag ggg aca	767
Gly Asp Thr Gly Phe Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr	
240 245 250 255	
ccg ggt tca cga gct gcc cac gtc ctc tcc agg aag gga ccc ggg tcc	815
Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro Gly Ser	
260 265 270	

acg	agc	tgc	cca	cgt	cct	ctc	cag	gag	ggg	aca	ccg	ggt	tca	cga	gct	863
Thr	Ser	Cys	Pro	Arg	Pro	Leu	Gln	Glu	Gly	Thr	Pro	Gly	Ser	Arg	Ala	
			275					280					285			
gcc	cac	gca	ctt	tcc	agg	aag	gga	ccc	cgg	gtt	cag	gtc	tcc	tgc	cgg	911
Ala	His	Ala	Leu	Ser	Arg	Lys	Gly	Pro	Arg	Val	Gln	Val	Ser	Cys	Arg	
		290					295					300				
ccc	aca	tcg	tgc	ctt	tgt	gta	aat	cag	aag	aaa	gat	gag	gaa	cag	gcc	959
Pro	Thr	Ser	Cys	Leu	Cys	Val	Asn	Gln	Lys	Lys	Asp	Glu	Glu	Gln	Ala	
		305				310					315					
ctc	ctc	tct	ctc	cag	gca	ggc	ttt	ggt	gga	ggg	gct	gga	tct	cct	gcc	1007
Leu	Leu	Ser	Leu	Gln	Ala	Gly	Phe	Gly	Gly	Gly	Ala	Gly	Ser	Pro	Ala	
					325					330					335	
gca	cct	tcc	ctg	gca	ggg	cac	cct	gtg	ctt	gag	ccc	cag	aac	tgc	agg	1055
Ala	Pro	Ser	Leu	Ala	Gly	His	Pro	Val	Leu	Glu	Pro	Gln	Asn	Cys	Arg	
				340					345					350		
cgg	ccg	gca	gag	aag	ggg	tcc	atg	atg	gcg	cct	cgg	tgc	gca	gcc	ttg	1103
Arg	Pro	Ala	Glu	Lys	Gly	Ser	Met	Met	Ala	Pro	Arg	Cys	Ala	Ala	Leu	
			355					360					365			
gac	ctg	ccc	cca	tgg	acc	tgg	gaa	cct	ccc	ggc	tct	tcc	cac	tcg	gga	1151
Asp	Leu	Pro	Pro	Trp	Thr	Trp	Glu	Pro	Pro	Gly	Ser	Ser	His	Ser	Gly	
		370					375					380				
aag	gaa	ggc	tct	ggg	cat	gga	ggt	cgg	cca	ggc	ccc	atc	ccc	gta	ccc	1199
Lys	Glu	Gly	Ser	Gly	His	Gly	Gly	Arg	Pro	Gly	Pro	Ile	Pro	Val	Pro	
		385				390					395					
tgg	ccc	ttc	ttc	ctg	ctt	cct	gtt	tgt	cac	tgc	ccc	ggg	gcc	ttt	gca	1247
Trp	Pro	Phe	Phe	Leu	Leu	Pro	Val	Cys	His	Cys	Pro	Gly	Ala	Phe	Ala	
	400				405					410					415	
cct	gca	ttc	cct	ctc	tct	gtg	agt	gtc	ctg	ggg	ccc	gtt	acc	cac	gtc	1295
Pro	Ala	Phe	Pro	Leu	Ser	Val	Ser	Val	Leu	Gly	Pro	Val	Thr	His	Val	
				420					425					430		
acc	gtc	cca	gga	tac	ctt	ttc	ttt	tct	ttc	tct	ctc	tcc	agc	ttt	att	1343
Thr	Val	Pro	Gly	Tyr	Leu	Phe	Phe	Ser	Phe	Ser	Leu	Ser	Ser	Phe	Ile	
			435					440					445			
gag	gta	tag	ttg	aca	attc	agg	acg	gtgt	gcact	caagg	tat	gcag	cat			1392
Glu	Val	*														
caca	acctga	cacac	gtagg	catt	gtg	aaa	tgagt	ccc	ac	aatt	ggg	ccta	atta	acac	ac	1452
ccat	cacctt	acat	ggttac	ttct	ttt	ctgt	ggtg	aga	aca	ctaa	atttta	aatag	agg	ac		1512
acac	gcctg	ggca	acatag	tgag	acc	ctg	tct	taca	aaa	tata	aaaaaaa	ttat	ctg	gac		1572
gtgg	tgggtgc	acac	ctgtgg	tccc	agct	ac	ttgg	gaag	ct	gagg	ctgg	gag	aat	ca	cttga	1632
gcct	gggagg	cgg	agg	ttgc												

<210> 11
<211> 449
<212> PRT
<213> Homo sapiens

<400> 11
Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala
1 5 10 15
Ala His Val Val Ser Arg Lys Gly Pro Gly Ser Thr Ser Cys Pro Arg
20 25 30
Pro Leu Gln Glu Arg Thr Arg Val His Glu Leu Ala Thr Ser Ser Ala
35 40 45
Gly Arg Asp Pro Gly Ser Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly
50 55 60
Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro Arg
65 70 75 80
Val His Glu Leu Pro Thr Ser Ser Pro Gly Arg Asp Pro Gly Ser Thr
85 90 95
Asn Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala
100 105 110
His Val Leu Ser Arg Arg Gly His Arg Val His Glu Leu Pro Thr Pro
115 120 125
Ser Pro Gly Arg Asp Pro Gly Phe Met Ser Cys Pro Arg Pro Leu Gln
130 135 140
Glu Gly Thr Arg Val His Glu Leu Pro Thr Pro Ser Pro Gly Gly Asp
145 150 155 160
Pro Gly Pro Arg Ala Ala His Val Val Asn Gly Lys Gly Pro Gly Ser
165 170 175
Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Arg Val His Glu Leu
180 185 190
Pro Thr Arg Ser Pro Gly Gly Asp Thr Gly Phe Thr Ser Cys Pro Arg
195 200 205
Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala His Val Leu Ser
210 215 220
Arg Arg Gly His Arg Val His Glu Leu Pro Thr Ser Ser Pro Gly Gly
225 230 235 240
Asp Thr Gly Phe Thr Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro
245 250 255
Gly Ser Arg Ala Ala His Val Leu Ser Arg Lys Gly Pro Gly Ser Thr
260 265 270
Ser Cys Pro Arg Pro Leu Gln Glu Gly Thr Pro Gly Ser Arg Ala Ala
275 280 285
His Ala Leu Ser Arg Lys Gly Pro Arg Val Gln Val Ser Cys Arg Pro
290 295 300
Thr Ser Cys Leu Cys Val Asn Gln Lys Lys Asp Glu Glu Gln Ala Leu
305 310 315 320
Leu Ser Leu Gln Ala Gly Phe Gly Gly Gly Ala Gly Ser Pro Ala Ala
325 330 335
Pro Ser Leu Ala Gly His Pro Val Leu Glu Pro Gln Asn Cys Arg Arg
340 345 350
Pro Ala Glu Lys Gly Ser Met Met Ala Pro Arg Cys Ala Ala Leu Asp
355 360 365
Leu Pro Pro Trp Thr Trp Glu Pro Pro Gly Ser Ser His Ser Gly Lys
370 375 380
Glu Gly Ser Gly His Gly Gly Arg Pro Gly Pro Ile Pro Val Pro Trp
385 390 395 400

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.20.438888>; this version posted April 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]